

SAIC0008-CON1
Serial No. 10/760,472

2

PATENT

AMENDMENTS TO THE CLAIMS

1.-32. (Cancelled)

33. (Currently amended) A system for automatically cataloguing documents located in multiple heterogeneous repositories to facilitate document search and retrieval capabilities, the system comprising:

a scanning tool for scanning the multiple heterogeneous repositories to collect keywords for the documents located therein;

an keyword index ~~[[to]]~~ of the documents built using the collected keywords;

a mapping tool for cataloguing the documents using the keyword index to one or more classes, each of the one or more classes including keywords representative of that class; and

a computing device for creating metadata ~~indicative of~~ for each of the catalogued documents and updating the index with the created metadata ~~indexing each of the documents in an index of an integrated library according to the metadata~~, wherein the created metadata for each of the catalogued documents indexed within the meta-index is stored in a pre-defined data structure including at least one of the following attributes a keyword, one or more matched words, and a classmark, further wherein the index retains characteristics of each of the multiple heterogeneous repositories as applied to each of the documents; and

a retrieving tool such that a user may access one or more of the documents within the multiple heterogeneous repositories corresponding to a user search request by utilizing the metadata in the pre-defined data structure index, and further wherein the characteristics of the multiple heterogeneous repositories are transparent to the user when one or more of the documents are accessed ~~using the index~~.

SAIC0008-CON1**3****PATENT**

Serial No. 10/760,472

34. (Cancelled)

35. (Cancelled)

36. (Previously presented) The system according to claim 33, wherein the metadata is stored in eXtensible Markup Language (XML) format.

37. (Previously presented) The system according to claim 33, wherein the metadata is stored in Resource Description Framework (RDF) format.

38. (Previously presented) The system according to claim 33, wherein the scanning tool is at least one spider.

39. (Previously presented) The system according to claim 33, wherein the mapping tool is a domain ontology.

40. (Previously presented) The system according to claim 39, wherein the domain ontology is a classification hierarchy.

41. (Previously presented) The system according to claim 33, wherein the mapping tool is a neural network.

US2000 10234096.1

SAIC0008-CON1

4

PATENT

Serial No. 10/760,472

42. (Currently amended) A method for automatically cataloguing documents located in multiple heterogeneous repositories to facilitate document search and retrieval capabilities, comprising:

scanning the multiple heterogeneous repositories to collect keywords from the documents located therein;

building an ~~keyword~~ index to the documents stored in the multiple heterogeneous repositories using the collected keywords;

cataloguing the documents using the ~~keyword~~ index into predetermined classes, wherein the cataloguing is performed using at least one mapping tool;

creating metadata information, including identification of the predetermined class, for each of the catalogued documents; and

updating the index with the created metadata ~~indexing each of the documents in an index of an integrated library according to the metadata~~, wherein the created metadata for each of the catalogued documents indexed within the index is stored in a pre-defined data structure including at least one of the following attributes a keyword, one or more matched words, and a classmark, further wherein the index retains the characteristics of each of the multiple heterogeneous repositories as applied to each of the documents; and

retrieving such that a user may access one or more of the documents within the multiple heterogeneous repositories corresponding to a user search request by utilizing the metadata in the pre-defined data structure index, and further wherein the characteristics of the multiple heterogeneous repositories are transparent to the user when one or more of the documents are accessed ~~using the index~~.

SAIC0008-CON1**5****PATENT**

Serial No. 10/760,472

43. (Previously presented) The method of claim 42, wherein scanning the at least one information repository to collect keywords is performed by a spider.

44. (Previously presented) The method of claim 42, wherein the metadata information is stored in the eXtensible Markup Language (XML) format.

45. (Previously presented) The method of claim 42, wherein the metadata information is stored in the Resource Description Framework (RDF) format.

46.-49. (Cancelled)